Kalpa Publications in Computing
Volume 22, 2025, Pages 271–278
Proceedings of The Sixth International Conference on Civil and Building Engineering Informatics

CB-YOLOv5: Streetlight Detection Based on Low-light Images in High-interference Environment

Shiqi Zhang[1], Jingyuan Tang[1], Jingke Hong[1]
[1] School of Management Science and Real Estate, Chongqing University, Chongqing, China

Abstract

The monitoring and operation maintenance (O&M) of urban streetlights is crucial for traffic safety and socio-economic development. However, how to accurately and robustly detect streetlights in low-light and high- interference environments is still a problem that concerns researchers. In recent years, deep learning has made remarkable progress in the field of object detection, among which the single-stage detection algorithm represented by You Only Look Once (YOLO) shows a satisfactory detection effect. It brings a new opportunity to detect streetlights based on images collected in a complicated street environment. Therefore, this study proposes an improved YOLOv5 model, as CB-YOLOv5, to accurately and robustly detect streetlights based on low-light images with high interferences. This proposed model integrates a Convolutional Block Attention Module (CBAM) and Bidirectional Feature Pyramid Network (BiFPN) to enhance its learning ability of spatial and channel dimension feature information, promote information fusion and transfer between multi-scale objects. Experimental results show that compared with the standard YOLOv5 algorithm, the proposed CB-YOLOv5 model can achieve significant improvement in accuracy and ability of interference resistance in streetlight detection tasks. The mAP0.5 reached 0.968, which is 23.5% higher than that of the standard YOLOv5 algorithm. In general, the CB-YOLOv5 model provides a new method to detect small objects in low-light and complex scenes. The developed method is also expected to provide a theoretical basis for automated monitoring and operation maintenance of urban lighting facilities.

1 Introduction

With the acceleration of urbanization, the effective management and maintenance of lighting infrastructure, as a key factor in urban safety, residents' quality of life and energy consumption, has become increasingly important. As a core component of the urban lighting system, streetlights are not only related to pedestrian and vehicle safety at night but also directly affect the energy efficiency and

I. Hkust, J. Cheng and Y. Yantao (eds.), ICCBEI 2025 (Kalpa Publications in Computing, vol. 22), pp. 271–278

environmental protection of the city. However, the traditional manual inspection method is inefficient and cannot achieve real-time monitoring and rapid response to the status of streetlights(Tosti et al., 2021). Therefore, the development of an efficient and accurate automatic streetlight recognition technology is of great significance for improving the intelligent level of urban management.

In recent years, the rapid development of deep learning technology, especially the progress in the field of target detection, has provided new solutions for streetlight recognition(T. Zhang, Dai, Song, Zhao, & Zhang, 2023). The YOLO series (You Only Look Once), as a leader in the field of real-time object detection, shows great potential in multiple application scenarios with its efficiency and accuracy(Y. Zhang et al., 2022). YOLOv5, as the latest generation, further improves detection speed and accuracy by optimizing the network structure and loss function(Wu, Wang, & Liu, 2021). However, in the face of complex and changeable urban environments, such as light change, uneven light, easy to be blocked, easy to confuse and other challenges, a single object detection model is often difficult to achieve the ideal recognition effect.

To overcome these challenges, in this study, the Convolutional Block Attention Module (CBAM) and Bidirectional Feature Pyramid Network (BiFPN) are introduced. CBAM enhances the feature representation through two dimensions of channel and space, enabling the model to focus on key information and ignore irrelevant background, so that the streetlight spots are not easy to be confused with other similar objects around (such as camera spots, Traffic light spots, etc.), thus improving the detection accuracy. BiFPN effectively solves the detection problem caused by the change of target scale through cross-scale feature fusion and enhances the detection ability of the model for small targets (such as distant streetlights). At the same time, combined with the efficient detection framework of YOLOv5, this study aims to build a high-performance streetlight recognition model to achieve accurate and rapid recognition of urban streetlights.

The main objectives of this study include: 1) For urban lighting infrastructure with streetlights as an example, a lightweight lighting facility recognition algorithm with both accuracy and stability is developed by using dashcams in unfavorable visual environments such as insufficient light, uneven light, shape occlusion and background change, and a streetlight recognition model based on CB-YOLOv5 is designed and implemented. Realize real-time troubleshooting of lighting facility fault points; 2) The recognition performance of the model under low-light and high-interference environments was verified by experiments; 3) Compare and analyze the baseline YOLOv5 model to evaluate the improvement effect and practical application value of the model.

2 Method

2.1 Network Architecture of CB-YOLOv5

In the YOLOv5 model, to enhance the accuracy and robustness of the model's recognition of streetlights in complex and changeable environments, especially in the case of high occlusion and high interference, we took the following two improvement measures: 1) The CBAM module was introduced in the Backbone part. We chose to embed CBAM in the Backbone part of YOLOv5. By combining channel attention and spatial attention, CBAM module helps the model focus on the more critical feature information for street light detection, thus improving the accuracy and robustness of detection. Especially in the complex and changeable environment, the introduction of CBAM makes the model more effective in dealing with various interference factors, such as lighting changes, occlusion, etc., to accurately recognize streetlights. 2) The BiFPN structure is used for the Neck part. To make full use of multi-scale feature information and enhance the model's adaptability to complex environments, we innovatively introduced the BiFPN structure into the Neck part of YOLOv5. BiFPN is known for its excellent feature fusion and transfer capabilities, which can help the Neck part more effectively integrate feature maps from different levels and scales. With the introduction of BiFPN, the model can more accurately capture and identify streetlights of different sizes, thus significantly improving the detection performance. Figure 1 is the overall framework of the proposed method.

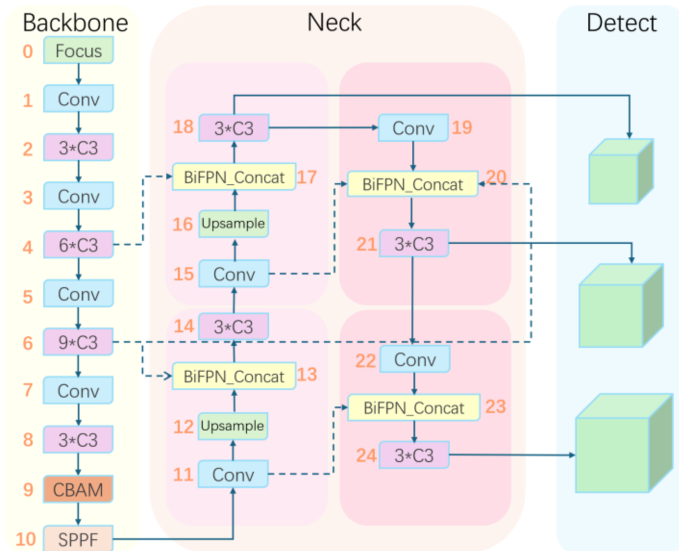

Figure 1: CB-YOLOv5 network structure

2.2 Bidirectional Feature Pyramid Network (BiFPN)

To solve the problem of multi-scale feature fusion in target detection tasks, we propose a novel Bidirectional Feature Pyramid Network (BiFPN). BiFPN's design was inspired by the bidirectional Feature Pyramid (FPN) and Path Aggregation Network (PANet), but has significantly improved upon it. The structure of BiFPN is shown in Figure 2.

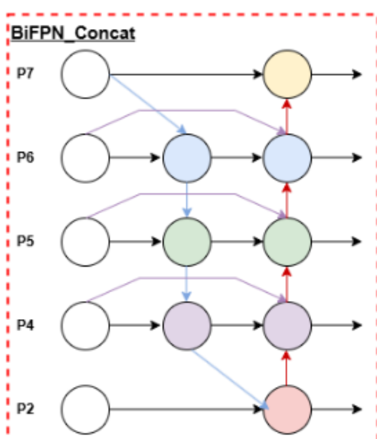

Figure 2: BiFPN structure

Bidirectional cross-scale connection

BiFPN retains the top-down and bottom-up paths in FPN and PANet for more comprehensive feature information transfer. This bidirectional connection allows the network to efficiently capture context information from different scales, thereby enhancing feature representation.

Weighted feature fusion

BiFPN uses weighted feature fusion to fuse input feature layers of different resolutions. These weights are trained by the neural network and used to adjust the importance of different features. In this way, the network can adaptively select and fuse the most relevant features, thus improving the effectiveness and accuracy of feature fusion.

Studies have shown that BiFPN can improve the efficiency and performance of models(M., R., & Q.V., 2020), especially for object detection, instance segmentation, and other related computer vision tasks(Wang, Pan, Lu, & Zhang, 2023). Since there may be various interference factors (such as light changes, shadows, occlusion, etc.) in street view in low-light environment, traditional recognition methods may be greatly affected. BiFPN, through its powerful feature fusion and weight allocation capabilities, can better deal with these interference factors and improve the robustness of identification.

2.3 Convolutional Block Attention Module (CBAM)

CBAM is a simple but effective attention module. It is designed to be flexible, easily integrated into various existing CNN architectures, and is widely used in a variety of visual recognition tasks. This dual attention mechanism enables the model to focus on key channels and spatial positions in input features. This adaptive re-calibration mechanism helps the model to focus on key features better and improve recognition performance(Zhu, Lyu, Wang, & Zhao, 2021). Experiments show that CBAM is integrated into different models, and the performance of the model is greatly improved on different classification and detection data sets, which proves the effectiveness of this module(Xudong, Shuai, & Chaoqun, 2022). In the street view we obtained, there are elements that are easy to confuse the streetlight spots, such as camera light spots. The occlusion of leaves and signs, as well as the fuzzy imaging caused by low-light environments, make the recognition task difficult. The CBAM attention module is introduced into the model to help CB-YOLOv5 resist confusing information and focus on useful target objects. The CBAM module is shown in Figure 3.

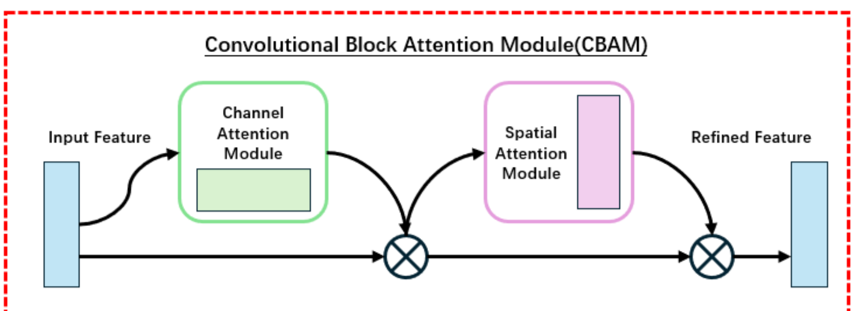

Figure 3: CBAM structure

3 Experiments

3.1 Datasets

The experimental dataset includes 200 images which were extracted by a video of a real streetlight monitoring task in Chongqing, China. This video was captured in low-light environment with complex backgrounds. The data was divided into a training set and a validation set in a 7:1 ratio.

3.2 Details

Our experiment was conducted under the Pytroch 2.0.0 framework of Python 3.11.9, using a personal computer with the NVIDIA GeForce RTX 4070 GPU for training and testing. The 640×640 RGB images were used as an input. In the training phase, we use part of the pre-trained model from yolov5s, because CB-YOLOv5 and YOLOv5 share some parts of the Backbone and the Neck, there are many weights that can be transferred from YOLOv5s to CB-YOLOv5, by using these weights we can save a lot of training time.

3.3 Evaluation Indicators

In this paper, precision rate, recall rate, average precision and mean average precision are used as evaluation indexes. The following equations define these indices:

$$P = \frac{TP}{TP + FP} \tag{1}$$

$$R = \frac{TP}{TP + FN} \tag{2}$$

$$AP = \int_0^1 P(R)dR \tag{3}$$

Where P, R, and AP represent precision rate, recall rate, and average precision, respectively. TP, FP, and FN define true positives, false positives, and false negatives. The precision rate reflects the proportion of the real number of targets predicted by the model in all detected targets, the calculation formula is formula (1). Recall rate reflects the proportion of the number of real targets predicted by the model in all real detection targets, the calculation formula is formula (2). Average precision reflects the overall detection precision of the model by calculating the area enclosed by the precision-recall (PR) curves, the calculation formula is formula (3). mAP (Mean Average Precision) is a common index used to measure the overall performance of an object detection model. It considers the Recall and Precision of the model. The higher the mAP value, the better the model performance. mAP0.5 indicates the mAP calculated when the IOU threshold is 0.5. This means that a prediction is considered correct when the overlap between the prediction box and the real box is 50% or more.

3.4 Results

We used the pre-trained weights of the YOLOv5s model on the COCO dataset. The learning rate was set to 0.01, and the Stochastic Gradient Descent (SGD) method was used to iterate. In the training period, batch size was set to 16 and run 300 epochs. The training results are shown in Figure 4.

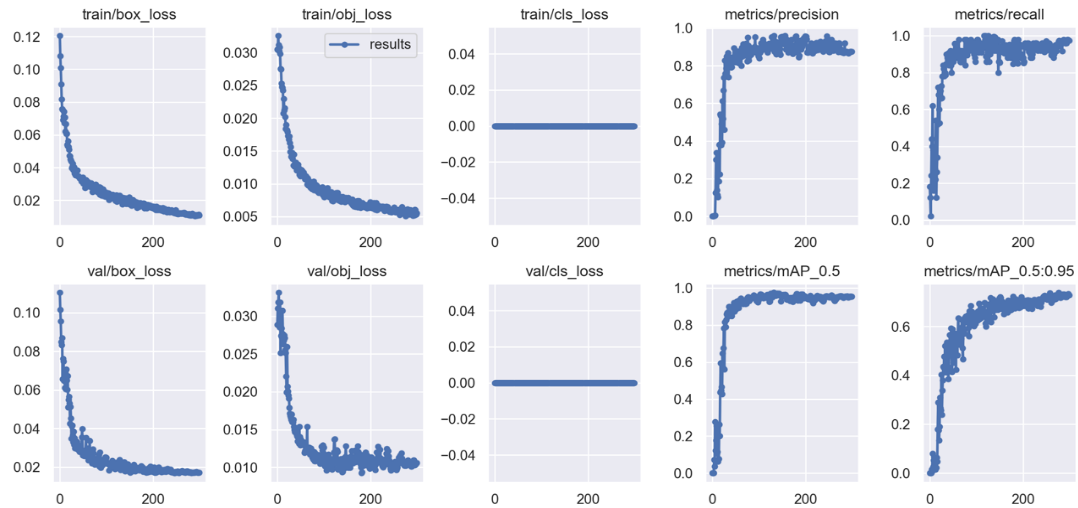

Figure 4: Training results of our model

The CB-YOLOv5 model has strong resistance to occlusion and high interference. It can be seen from the test results of some streetlights in Figure 5 that even the streetlights shielded by leaves with only a small part of the light spots exposed can be identified, the camera spots very similar to the streetlights can be excluded, and the white light spots on the surrounding buildings can also be avoided. At the same time, in the low-light environment, due to the existence of the CBAM attention mechanism, the CB-YOLOv5 model can also overcome the influence of the harsh environment and focus on the characteristics of streetlights, to achieve accurate recognition.

Figure 5: Some streetlight detection results in complex scenarios

Comparative tests and evaluations were conducted on different network structures, and the results are shown in Table 1. As shown in Table 1, adding a CBAM module to the backbone for standard YOLOv5s can increase mAP0.5 by 16.3%. Based on adding the CBAM module, the BiFPN structure can further improve mAP0.5 by 7.2%.

Models	$mAP_{0.5}$	P	R
YOLOv5s	0.733	0.713	0.709
YOLOv5s+CBAM	0.896	0.898	0.814
YOLOv5s+CBAM+BiFPN	0.968	0.921	0.962

Table 1: Detection results of different models

4 Conclusions

In this paper, we proposed an improved YOLOv5 model to improve the efficiency and performance of streetlight recognition monitoring by combining CBAM and BiFPN. By utilizing CBAM, the loss in feature extraction is reduced, ensuring that more relevant and distinct features are retained. In addition, BiFPN facilitates multi-scale feature fusion and propagation, which is critical for accurately detecting streetlights of different sizes and distances.

Results obtained on our customized dataset show a significant improvement in the detection accuracy of the CB-YOLOv5 model, with mAP0.5 exceeding the baseline YOLOv5 model by 23.5%. In this study, the model will be applied to streets with different lighting conditions and road environments to verify its practicability and generalization ability. In addition, the CB-YOLOv5 model will be further applied to the recognition of streetlight poles in future research, and then the recognition performance of the model will be verified again, to generalize it for the recognition of other small objects under complex backgrounds and harsh environment conditions. Based on the two identification results of the streetlight spot and the streetlight pole, the fault point of the lighting facility is investigated, and the brightness threshold can be set subsequently to classify the working performance of the streetlight with high detail.

In addition, this study will continue to explore the model lightweight strategy in the future, in order to ensure recognition accuracy, reduce the consumption of computing resources, and facilitate deployment on edge devices.

Acknowledgements

The authors are grateful for the Fundamental Research Funds for the Central Universities of Ministry of Education of China (No. 2023CDJKYJH023).

References

M., T., R., P., & Q.V., L. (2020). EfficientDet: Scalable and efficient object detection %J Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 10778-10787.

Tosti, F., Gagliardi, V., Ciampoli, L. B., Benedetto, A., Threader, S., & Alani, A. M. (2021). *Integration of Remote Sensing and Ground-Based Non-Destructive Methods in Transport*

Infrastructure Monitoring: Advances, Challenges and Perspectives. Paper presented at the 2021 IEEE Asia-Pacific Conference on Geoscience, Electronics and Remote Sensing Technology (AGERS).

Wang, J., Pan, Q., Lu, D., & Zhang, Y. (2023). An Efficient Ship-Detection Algorithm Based on the Improved YOLOv5. *Electronics*, *12*(17). doi:10.3390/electronics12173600

Wu, T.-H., Wang, T.-W., & Liu, Y.-Q. (2021). *Real-Time Vehicle and Distance Detection Based on Improved Yolo v5 Network*. Paper presented at the 2021 3rd World Symposium on Artificial Intelligence (WSAI).

Xudong, D., Shuai, Y., & Chaoqun, D. (2022). A lightweight vehicles detection network model based on YOLOv5 %J Engineering Applications of Artificial Intelligence. *113*.

Zhang, T., Dai, J., Song, W., Zhao, R., & Zhang, B. (2023). OSLPNet: A neural network model for street lamp post extraction from street view imagery. *Expert Systems with Applications*, *231*. doi:10.1016/j.eswa.2023.120764

Zhang, Y., Guo, Z., Wu, J., Tian, Y., Tang, H., & Guo, X. (2022). Real-Time Vehicle Detection Based on Improved YOLO v5. *Sustainability, 14*(19). doi:10.3390/su141912274

Zhu, X., Lyu, S., Wang, X., & Zhao, Q. (2021). *TPH-YOLOv5: Improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios.* Paper presented at the Proceedings of the IEEE/CVF international conference on computer vision.